



Related Structures

Sequences producing significant alignments:	Score (bits)	E Value	
gi 626062 pir JC2217 major surface glycoprotein 5 - Pneumo...	32	1.9	
gi 309818 gb AAA33784.1 antigen	32	1.9	
gi 22958582 ref ZP_00006250.1 COG1360: Flagellar motor pro...	29	11	
gi 38075587 ref XP_356708.1 similar to transcription facto...	29	15	L
gi 23600471 gb AAN39050.1 putative phosphate transporter O...	29	15	
gi 7341204 gb AAF61213.1 NS protein [pig-tailed macaque pa...	28	20	
gi 20150320 pdb 1J4T A Chain A, Structure Of Artocarpin: A ...	28	20	S
gi 22122697 ref NP_666274.1 hypothetical protein MGC38046 ...	28	20	L
gi 20150316 pdb 1J4S A Chain A, Structure Of Artocarpin: A ...	28	20	S
gi 26347287 dbj BAC37292.1 unnamed protein product [Mus mu...	28	20	L
gi 50255367 gb EAL18102.1 hypothetical protein CNBK1230 [C...	28	20	
gi 34873522 ref XP_220904.2 similar to hypothetical protei...	28	27	L
gi 309820 gb AAA33785.1 Antigen	28	27	
gi 16041188 dbj BAB69761.1 hypothetical protein [Macaca fa...	28	27	
gi 34899222 ref NP_910957.1 alcohol dehydrogenase-like pro...	28	27	
gi 48845752 ref ZP_00300025.1 COG3209: Rhs family protein ...	27	36	
gi 15226565 ref NP_179743.1 protein kinase family protein ...	27	36	
gi 2134685 pir I53411 acetylcholinesterase (EC 3.1.1.7), m...	27	36	L

gi 25151456 ref NP_741140.1 	prion-like Q/N-rich domain pro...	27	36	L
gi 30089741 gb AAP20845.1 	hypothetical protein [Oryza sati...	27	36	
gi 499385 emb CAA53511.1 	collectin-43 [Bos taurus]	27	36	L
gi 50355694 ref NP_001002237.1 	collectin-43 [Bos taurus] >...	27	36	L
gi 1083017 pir A53570	collectin-43 - bovine	27	36	L
gi 45546404 ref ZP_00186489.1 	COG0642: Signal transduction...	27	48	
gi 38086162 ref XP_133316.2 	similar to hypothetical protei...	27	48	L
gi 20808945 ref NP_624116.1 	hypothetical protein [Thermoan...	27	48	
gi 6753710 ref NP_034228.1 	opsin (encephalopsin); encephal...	27	48	L
gi 45547389 ref ZP_00187440.1 	COG0642: Signal transduction...	27	48	
gi 22971781 ref ZP_00018707.1 	hypothetical protein [Chloro...	27	48	
gi 1200129 emb CAA88906.1 	alpha-adducin, hypertensive phen...	27	65	L
gi 30794342 ref NP_851369.1 	surfactant, pulmonary-associat...	27	65	L
gi 6851286 gb AAF29504.1 	alpha adducin [Mus musculus]	27	65	L
gi 21166364 gb AAM43814.1 	PCM1 [Takifugu rubripes]	27	65	
gi 423283 pir S33603	surfactant protein D - bovine	27	65	
gi 12643483 sp Q63028 ADDA_RAT	Alpha adducin (Erythrocyte a...	27	65	L
gi 23008382 ref ZP_00049851.1 	COG4456: Virulence-associate...	27	65	
gi 31982388 ref NP_783630.2 	conglutinin 1 [Bos taurus] >gi...	27	65	L
gi 47216258 emb CAG05954.1 	unnamed protein product [Tetrao...	27	65	
gi 10719868 sp Q9QYC0 ADDA_MOUSE	Alpha adducin (Erythrocyte...	27	65	L
gi 27676688 ref XP_218355.1 	similar to hypothetical protei...	27	65	L
gi 34873860 ref XP_343971.1 	similar to Mox-1 [Rattus norve...	27	65	L
gi 27711226 ref XP_213841.1 	similar to RIKEN cDNA 1110002H...	27	65	L
gi 47211747 emb CAF94312.1 	unnamed protein product [Tetrao...	27	65	
gi 21706416 gb AAH34368.1 	Add1 protein [Mus musculus]	27	65	L
gi 12844579 dbj BAB26417.1 	unnamed protein product [Mus mu...	27	65	L
gi 7304861 ref NP_038485.1 	adducin 1 (alpha) [Mus musculus...	27	65	L
gi 461774 sp P23805 CONG_BOVIN	Conglutinin precursor >gi 34...	27	65	L
gi 8850211 ref NP_058686.1 	adducin 1, alpha; alpha-ADD [Ra...	27	65	L
gi 26988441 ref NP_743866.1 	fumarylacetoacetate hydrolase ...	27	65	
gi 29570599 emb CAD69922.1 	surfactant protein D [Bos taurus]	27	65	L
gi 14579673 gb AAK69357.1 	pericentriolar material 1 protei...	27	65	
gi 49258206 ref NP_001001856.1 	collectin 46 [Bos taurus] >...	27	65	L
gi 395268 emb CAA50665.1 	conglutinin [Bos taurus]	27	65	L
gi 6754680 ref NP_034921.1 	mesenchyme homeobox 1 [Mus musc...	27	65	L
gi 1083589 pir S54147	alpha adducin - rat	27	65	
gi 47216257 emb CAG05953.1 	unnamed protein product [Tetrao...	27	65	
gi 50510951 dbj BAD32461.1 	mKIAA1482 protein [Mus musculus]	26	87	L
gi 12842851 dbj BAB25758.1 	unnamed protein product [Mus mu...	26	87	L
gi 34881239 ref XP_228551.2 	similar to CCG1 [Rattus norveg...	26	87	L
gi 18959276 ref NP_036316.1 	forkhead box D4-like 1 [Homo s...	26	87	L
gi 32405518 ref XP_323372.1 	predicted protein [Neurospora ...	26	87	
gi 45190590 ref NP_984844.1 	AEL017Wp [Eremothecium gossypi...	26	87	
gi 34451549 gb AAQ72339.1 	FOXD4 [Gorilla gorilla]	26	87	
gi 39979247 emb CAE85616.1 	hypothetical protein [Neurospor...	26	87	
gi 45361285 ref NP_989220.1 	hypothetical protein MGC75588 ...	26	87	L
gi 13929054 ref NP_113939.1 	Shank1; GKAP/SAPAP interacting...	26	116	L
gi 47215010 emb CAG03150.1 	unnamed protein product [Tetrao...	26	116	
gi 48256849 gb AAT41626.1 	collagen type IX-like [Ciona int...	26	116	

gi 50513971 pdb 1VBP A Chain A, Crystal Structure Of Artoca...	26	116	S
gi 22001985 sp Q9WV48 SHK1 RAT SH3 and multiple ankyrin rep...	26	116	L
gi 38086600 ref XP_145592.3 similar to synaptic SAPAP-inte...	26	116	L
gi 50513963 pdb 1VBO A Chain A, Crystal Structure Of Artoca...	26	116	S
gi 27378759 ref NP_770288.1 bll3648 [Bradyrhizobium japoni...	26	116	
gi 7516461 pir E72614 hypothetical protein APE1374 - Aerop...	26	116	
gi 7025451 gb AAF35887.1 somatostatin receptor-interacting...	26	116	L
gi 4850168 gb AAD04569.2 synaptic SAPAP-interacting protei...	26	116	L
gi 11968152 ref NP_057232.1 SH3 and multiple ankyrin repea...	26	116	L
gi 7519945 pir A58801 mannose-specific lectin KM+ - Artoca...	26	116	
gi 26331236 dbj BAC29348.1 unnamed protein product [Mus mu...	25	156	L
gi 2739210 gb AAB94628.1 Eph-like receptor tyrosine kinase...	25	156	L
gi 50254791 gb EAL17536.1 hypothetical protein CNBM1020 [C...	25	156	
gi 46255173 ref YP_006085.1 putative hydrolase [Thermus th...	25	156	
gi 16799538 ref NP_469806.1 lin0462 [Listeria innocua Clip...	25	156	
gi 26245405 gb AAN77376.1 BTB/POZ and zinc-finger domain c...	25	156	L
gi 40254274 ref NP_775623.2 Eph receptor B1; ELK homolog [...]	25	156	L
gi 26331180 dbj BAC29320.1 unnamed protein product [Mus mu...	25	156	L
gi 9625581 ref NP_039832.1 Hypothetical protein HHV4gp016 ...	25	156	
gi 4758284 ref NP_004432.1 ephrin receptor EphB1 precursor...	25	156	L
gi 34870072 ref XP_213685.2 similar to CG33130-PA [Rattus ...]	25	156	L
gi 38089043 ref XP_133997.4 RIKEN cDNA 2310057J16 [Mus mus...	25	156	L
gi 38101266 gb EAA48254.1 hypothetical protein MG10317.4 [...]	25	156	
gi 21224968 ref NP_630747.1 conserved hypothetical protein...	25	156	
gi 4104411 gb AAD02030.1 Eph-like receptor tyrosine kinase...	25	156	
gi 49257790 gb AAH74709.1 Unknown (protein for MGC:69445) ...	25	156	
gi 34327988 dbj BAB67820.2 KIAA1927 protein [Homo sapiens]	25	156	L
gi 29179544 gb AAH48787.1 Similar to RIKEN cDNA 4930541M15...	25	156	
gi 16121180 ref NP_404493.1 hypothetical protein [Yersinia...	25	156	
gi 12328471 dbj BAB21131.1 P0416D03.19 [Oryza sativa (japo...	25	156	
gi 27721289 ref XP_217250.1 similar to Ephrin type-B recep...	25	156	L
gi 23613524 ref NP_704545.1 erythrocyte membrane protein 1...	25	156	

Alignments

☐ >gi|626062|pir||JC2217 major surface glycoprotein 5 - Pneumocystis carinii
 gi|425785|dbj|BAA04851.1| major surface glycoprotein [Pneumocystis carinii]
 gi|743315|prf||2012229A major surface glycoprotein:ISOTYPE=MSG5
 Length = 1076

Score = 31.6 bits (67), Expect = 1.9

Identities = 13/20 (65%), Positives = 13/20 (65%), Gaps = 3/20 (15%)

Query: 3 QGPAGSGWEEGSGSPPGVTP 22
 Q PAGS GSGSPP V P
 Sbjct: 816 QAPAGSS---GSGSPPAVPP 832

☐ >gi|309818|gb|AAA33784.1| antigen
Length = 386

Score = 31.6 bits (67), Expect = 1.9
Identities = 13/20 (65%), Positives = 13/20 (65%), Gaps = 3/20 (15%)

Query: 3 QGPAGSGWEEGSGSPPGVTP 22
Q PAGS GSGSPP V P
Sbjct: 126 QAPAGSS---GSGSPPAVPP 142

☐ >gi|22958582|ref|ZP_00006250.1| COG1360: Flagellar motor protein [Rhodobacter sp]
Length = 366

Score = 29.1 bits (61), Expect = 11
Identities = 11/16 (68%), Positives = 13/16 (81%), Gaps = 2/16 (12%)

Query: 4 GPAGSGWEEGSGSPPG 19
GPAG G EG+G+PPG
Sbjct: 189 GPAGPG--EGTGAPPG 202

☐ >gi|38075587|ref|XP_356708.1| ☒ similar to transcription factor Oct-3, long spli
[Mus musculus]
Length = 153

Score = 28.6 bits (60), Expect = 15
Identities = 11/17 (64%), Positives = 12/17 (70%), Gaps = 3/17 (17%)

Query: 2 MQ--GPAGSGWEEGSGS 16
MQ GP G+GWEE S S
Sbjct: 1 MQIGGP-GAGWEENSSS 16

☐ >gi|23600471|gb|AAN39050.1| putative phosphate transporter OsPT9 [Oryza sativa (cultivar-group)]
Length = 582

Score = 28.6 bits (60), Expect = 15
Identities = 12/18 (66%), Positives = 14/18 (77%), Gaps = 2/18 (11%)

Query: 5 PAGSGWEEGSGSPPGVTP 22
P+GS +GSGS PGVTP
Sbjct: 55 PSGSV--DGS GSGPGVTP 70

☐ >gi|7341204|gb|AAF61213.1| NS protein [pig-tailed macaque parvovirus]
Length = 672

Score = 28.2 bits (59), Expect = 20
Identities = 13/20 (65%), Positives = 13/20 (65%), Gaps = 3/20 (15%)

Query: 10 WE--EGSGSPP-GVTPLFSP 26
WE E SGSP TPL SP

Sbjct: 542 WESSEDSGSPPRSSTPLASP 561

☐ >[gi|20150320|pdb|1J4T|A](#) **S** Chain A, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150321|pdb|1J4T|B](#) **S** Chain B, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150322|pdb|1J4T|C](#) **S** Chain C, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150323|pdb|1J4T|D](#) **S** Chain D, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150324|pdb|1J4T|E](#) **S** Chain E, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150325|pdb|1J4T|F](#) **S** Chain F, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150326|pdb|1J4T|G](#) **S** Chain G, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150327|pdb|1J4T|H](#) **S** Chain H, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 2)
[gi|20150328|pdb|1J4U|A](#) **S** Chain A, Structure Of Artocarpin Complexed With Me-Alpha
[gi|20150329|pdb|1J4U|B](#) **S** Chain B, Structure Of Artocarpin Complexed With Me-Alpha
[gi|20150330|pdb|1J4U|C](#) **S** Chain C, Structure Of Artocarpin Complexed With Me-Alpha
[gi|20150331|pdb|1J4U|D](#) **S** Chain D, Structure Of Artocarpin Complexed With Me-Alpha
 Length = 149

Score = 28.2 bits (59), Expect = 20
 Identities = 8/11 (72%), Positives = 10/11 (90%)

Query: 4 GPAGSGWEEGS 14
 GP G+GW+EGS
 Sbjct: 12 GPGGNGWDEGS 22

☐ >[gi|22122697|ref|NP_666274.1|](#) **L** hypothetical protein MGC38046 [Mus musculus]
[gi|19343896|gb|AAH25600.1|](#) **L** Hypothetical protein MGC38046 [Mus musculus]
 Length = 280

Score = 28.2 bits (59), Expect = 20
 Identities = 16/28 (57%), Positives = 17/28 (60%), Gaps = 9/28 (32%)

Query: 6 AGSGWE-EGSGS-----PPGVTPLEFSP 26
 AGSG E EGS + PPG TP FSP
 Sbjct: 34 AGSG-EAEGSSASSPSLPPPG-TPAFSP 59

☐ >[gi|20150316|pdb|1J4S|A](#) **S** Chain A, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 1)
[gi|20150317|pdb|1J4S|B](#) **S** Chain B, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 1)
[gi|20150318|pdb|1J4S|C](#) **S** Chain C, Structure Of Artocarpin: A Lectin With Mannose Specificity (Form 1)

gi|20150319|pdb|1J4S|D **S** Chain D, Structure Of Artocarpin: A Lectin With Mannose
Specificity (Form 1)
Length = 149

Score = 28.2 bits (59), Expect = 20
Identities = 8/11 (72%), Positives = 10/11 (90%)

Query: 4 GPAGSGWEEGS 14
GP G+GW+EGS
Sbjct: 12 GPGGNGWDEGS 22

>gi|26347287|dbj|BAC37292.1| **L** unnamed protein product [Mus musculus]
Length = 280

Score = 28.2 bits (59), Expect = 20
Identities = 16/28 (57%), Positives = 17/28 (60%), Gaps = 9/28 (32%)

Query: 6 AGSGWE-EGSGS-----PPGVTPLFSP 26
AGSG E EGS + PPG TP FSP
Sbjct: 34 AGSG-EAEGSSASSPSLPPPG-TPAFSP 59

>gi|50255367|gb|EAL18102.1| hypothetical protein CNBK1230 [Cryptococcus neoforma
neoformans B-3501A]
Length = 598

Score = 28.2 bits (59), Expect = 20
Identities = 13/29 (44%), Positives = 13/29 (44%), Gaps = 14/29 (48%)

Query: 1 GMQGPA-----GSGW-----EEGSG 15
GMQGPA G GW EEG G
Sbjct: 518 GMQGPAGVMAGTGDVGMGWGGIQKEEGGG 546

>gi|34873522|ref|XP_220904.2| **L** similar to hypothetical protein 4932418K24 [Ratt
Length = 747

Score = 27.8 bits (58), Expect = 27
Identities = 8/11 (72%), Positives = 10/11 (90%)

Query: 16 SPPGVTPLFSP 26
SPPG +PLF+P
Sbjct: 112 SPPGASPLFTP 122

>gi|309820|gb|AAA33785.1| Antigen
Length = 417

Score = 27.8 bits (58), Expect = 27
Identities = 11/16 (68%), Positives = 11/16 (68%), Gaps = 3/16 (18%)

Query: 3 QGPAGSGWEEGSGSPP 18
Q PAGS GSGSPP

Sbjct: 160 QAPAGSS---GSGSPP 172

☐ >gi|16041188|dbj|BAB69761.1| hypothetical protein [Macaca fascicularis]
Length = 135

Score = 27.8 bits (58), Expect = 27
Identities = 14/28 (50%), Positives = 14/28 (50%), Gaps = 9/28 (32%)

Query: 4 GPAG-SGWEEGSGSP-----PGVTP 22
GP G SGWE G SP P VTP
Sbjct: 83 GPVGASGWEAGCASPQPTSLTPYPRVTP 110

☐ >gi|34899222|ref|NP_910957.1| alcohol dehydrogenase-like protein~contains EST C7
[Oryza sativa (japonica cultivar-group)]
Length = 462

Score = 27.8 bits (58), Expect = 27
Identities = 12/19 (63%), Positives = 12/19 (63%), Gaps = 2/19 (10%)

Query: 5 PAGSGWEEGSGSPPGVTP 23
PAGSG E GSPP V L
Sbjct: 405 PAGSGRGE--GSPPAVKSL 421

☐ >gi|48845752|ref|ZP_00300025.1| COG3209: Rhs family protein [Geobacter metallire]
Length = 2294

Score = 27.4 bits (57), Expect = 36
Identities = 12/19 (63%), Positives = 13/19 (68%), Gaps = 2/19 (10%)

Query: 4 GPAGSGWEEGSGSPPGVTP 22
GPAG GW E G+P VTP
Sbjct: 214 GPAGPGWLEVQGTP--VTP 230

Score = 21.8 bits (44), Expect = 1638
Identities = 8/11 (72%), Positives = 9/11 (81%), Gaps = 1/11 (9%)

Query: 12 EGS-GSPPGVT 21
+GS SPPGVT
Sbjct: 1376 DGSFASPPGVT 1386

☐ >gi|15226565|ref|NP_179743.1| protein kinase family protein [Arabidopsis thaliana]
gi|25387138|pir||G84601 probable protein kinase [imported] - Arabidopsis thaliana
gi|4567279|gb|AAD23692.1| putative protein kinase [Arabidopsis thaliana]
Length = 871

Score = 27.4 bits (57), Expect = 36
Identities = 8/10 (80%), Positives = 9/10 (90%)

Query: 17 PPGVTPLFSP 26

PPGVTPL +P
Sbjct: 286 PPGVTPLIAP 295

☐ >gi|2134685|pir||I53411 acetylcholinesterase (EC 3.1.1.7), minor splice form - h
(fragment)
gi|545225|gb|AAC60618.1| ☒ acetylcholinesterase; acetylcholine acetylhydrolase; A
sapiens]
Length = 72

Score = 27.4 bits (57), Expect = 36
Identities = 8/8 (100%), Positives = 8/8 (100%)

Query: 1 GMQGPAGS 8
GMQGPAGS
Sbjct: 4 GMQGPAGS 11

☐ >gi|25151456|ref|NP_741140.1| ☒ prion-like Q/N-rich domain protein PQN-25, Prion
domain protein (pqn-25) [Caenorhabditis elegans]
gi|16604137|gb|AAL27237.1| Prion-like-(q/n-rich)-domain-bearing protein protein 2
[Caenorhabditis elegans]
Length = 672

Score = 27.4 bits (57), Expect = 36
Identities = 8/9 (88%), Positives = 9/9 (100%)

Query: 17 PPGVTPLFS 25
PPG+TPLFS
Sbjct: 51 PPGLTPLFS 59

☐ >gi|30089741|gb|AAP20845.1| hypothetical protein [Oryza sativa (japonica cultivar
Length = 50

Score = 27.4 bits (57), Expect = 36
Identities = 8/10 (80%), Positives = 9/10 (90%)

Query: 6 AGSGWEEGSG 15
AG+GWEEG G
Sbjct: 26 AGAGWEEGGG 35

☐ >gi|499385|emb|CAA53511.1| ☒ collectin-43 [Bos taurus]
Length = 301

Score = 27.4 bits (57), Expect = 36
Identities = 15/26 (57%), Positives = 16/26 (61%), Gaps = 6/26 (23%)

Query: 1 GMQGPAG----SGWEEGSGSPPGVTP 22
GM GPAG SG +GS PPG TP
Sbjct: 50 GMPGPAGREGPSGR-QGSMGPPG-TP 73

☐ >gi|50355694|ref|NP_001002237.1| **L** collectin-43 [Bos taurus]
gi|27923966|sp|P42916|CL43 BOVIN **L** Collectin-43 precursor (CL-43) (43 kDa collect
gi|18252109|gb|AAL61855.1| **L** 43kDa collectin precursor [Bos taurus]
gi|18252111|gb|AAL61856.1| **L** 43kDa collectin precursor [Bos taurus]
 Length = 321

Score = 27.4 bits (57), Expect = 36
 Identities = 15/26 (57%), Positives = 16/26 (61%), Gaps = 6/26 (23%)

Query: 1 GMQGPAG----SGWEEGSGSPPGVTP 22
 GM GPAG SG +GS PPG TP
 Sbjct: 70 GMPGPAGREGPSGR-QGSMGPPG-TP 93

☐ >gi|1083017|pir||A53570 **L** collectin-43 - bovine
 Length = 301

Score = 27.4 bits (57), Expect = 36
 Identities = 15/26 (57%), Positives = 16/26 (61%), Gaps = 6/26 (23%)

Query: 1 GMQGPAG----SGWEEGSGSPPGVTP 22
 GM GPAG SG +GS PPG TP
 Sbjct: 50 GMPGPAGREGPSGR-QGSMGPPG-TP 73

☐ >gi|45546404|ref|ZP_00186489.1| COG0642: Signal transduction histidine kinase [R
 xylanophilus DSM 9941]
 Length = 1280

Score = 26.9 bits (56), Expect = 48
 Identities = 9/13 (69%), Positives = 10/13 (76%)

Query: 5 PAGSGWEEGSGSP 17
 P+ SG EEG GSP
 Sbjct: 1023 PSSSGREEGGGSP 1035

☐ >gi|38086162|ref|XP_133316.2| **L** similar to hypothetical protein [Mus musculus]
 Length = 616

Score = 26.9 bits (56), Expect = 48
 Identities = 11/20 (55%), Positives = 12/20 (60%), Gaps = 4/20 (20%)

Query: 7 GSGWEEGSG----SPPGVTP 22
 GSGWEEG G + P V P
 Sbjct: 521 GSGWEEGEGRTTSTEPSVIP 540

☐ >gi|20808945|ref|NP_624116.1| hypothetical protein [Thermoanaerobacter tengconge
gi|20517608|gb|AAM25720.1| hypothetical protein TTE2596 [Thermoanaerobacter tengc
 MB4]
 Length = 68